Importance of Context Awareness in NLP

Nour Matta, Nada Matta and Philippe Herr
LIST3N, University of Technology of Troyes, 12 Rue Marie Curie, 42060 10004 Troyes Cedex, France

Keywords: Context Awareness, NLP, Text Ambiguity, Knowledge.

Abstract: Context is a complex notion, that enables the understanding of happenings and concepts in an environment and the analysis of their influence (Adomavicius et al, 2011) As previously mentioned, context plays a major role in assigning meanings to words, sentences, and texts when dealing with text analysis. Multiple natural language processing approaches aim to consider "context" in analyzing the information extracted and applying a sort of word sense disambiguation (Adhikari et al, 2019). Numerous intelligence systems require knowledge of happening and are context dependent, but the definition of context and context elements used varies from one application to another based on needs. Context plays several roles in text analysis especially to reduce ambiguity and semantic extraction. In this paper, main influence of context on TextMining and NLP are shown.

1 INTRODUCTION

From a semantic perspective, if the purpose was identifying business events, the verb "fire" in the sentence "a company fired 50 employees" can be considered as a business event. In contrast, the same verb in "he fired the gun" is not a business event of interest. We can highlight in this example the necessity of the words' context in understanding their meaning. But when dealing with decisions, the impact of the information extracted must also be considered in the process for two main reasons. First, the same information can have a different impact on different entities. The information "a company fired 50 employees" is considered relatively important for the company's competitors or main clients since this information may be viewed as a sign of struggle in the company of interest but can be irrelevant for unrelated organizations. Second, the information can have a different impact based on the entities involved in the event. For instance, if the company firing employees is a small company of 60 employees in total, this event may mean that the company is more likely to be closing. But if the company originally had more than 5000 employees, firing 50 employees is more or less irrelevant to the financial state of the company.

The dependency on the context of the information while extracting knowledge, from the activity domain to the participating entities, the events mentioned, the time factor, and so on, must be considered in the analysis process. Furthermore, when extracting information from texts, knowledge representation is a required task to enable the accessibility, reuse, and learning process. When dealing with the development of strategies and decision-making based on information extracted from text, the context of this information must be considered to enable the understanding of the information along with the analysis of the importance and the impact of this information.

So, our main research question is: How to deal with the context-dependency of the words semantic in texts?

In this paper, the importance of context awareness is emphasized to consider Natural Language processing techniques.

2 CONTEXT AWARENESS

Bazire and Brézillon (Bazire et al, 2005) used 150 definitions from different domains such as computer science, philosophy, economy, and business, and tried to combine all and abstract key elements. Their research showed that context may be defined by six main components: (1) the constraints, (2) the influence, and (3) the behavior of (4) a system with specific tasks to implement, where the system can be a user or a computer. The context can also be

Matta, N., Matta, N. and Herr, P.
Importance of Context Awareness in NLP.
DOI: 10.5220/0012994700003838

In *Proceedings of the 16th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2024) - Volume 3: KMIS*, pages 280-286
ISBN: 978-989-758-716-0; ISSN: 2184-3228

categorized by its (5) nature and (6) structure. Figure 2.1 shows a representation of the context as the elements that define the context and the interactions or influences between the entities. Five elements are represented in 0 (1) the context, (2) the system, (3) the item, (4) the environment, and (5) the observer. The context is the overall group of entities and how they influence each other. The system is the machine or person having specific tasks to implement on an object in an environment. The item is the object that undergoes changes while the environment can be the organization, the location, or the time in which changes are happening. Finally, the observer is an external element that will have a different opinion on the happenings considering the context. The observer enables the consideration of different cultural backgrounds and social views that might affect the reaction, or the decision made facing an event (Matsumoto, 2007). We can also notice in 0 the interaction between the different entities and the context influence of the context (orange arrow) that each entity has.

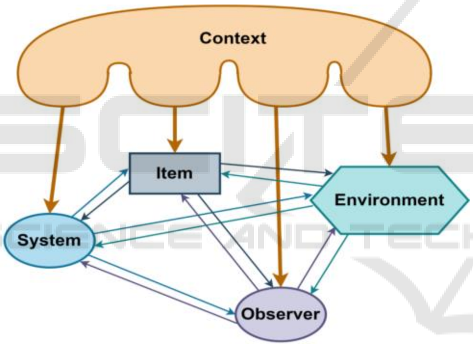


Figure 1: Context element definition and interaction.

3 CONTEXT IN TEXTMINING

3.1 TextMining and NLP

Text mining is the branch of artificial intelligence that aims to extract knowledge from structured and unstructured text (5). Text Mining enables the extraction of the knowledge available in stored textual data. Text Mining is mainly based on two components (5), the data mining and machine learning component and the computational linguistics also known as Natural Language Processing (NLP). On the data processing side, three phases can be identified (6). First, Information Retrieval allows the selection of documents of interest that are most likely related to the topic of interest. Second, data mining, machine learning algorithms, and probabilistic approaches allow the identification of patterns within the extracted data. Third, the Knowledge representation part allows the information extracted to be represented in a formal structure. On the other hand, NLP is used to simulate human's 'natural' understanding of languages to process textual data (Du, 2007). NLP researchers were first split into two divisions: stochastic and symbolic. Stochastic NLP consisted of probabilistic and statistical approaches, focusing on pattern recognition between texts. On the other hand, symbolic NLP also known as rule-based NLP was oriented on formal languages and generating syntax. NLP considers all linguistic levels (Liddy, 2001):

1. Phonetics is the study of the production of sounds
2. Phonology is the study of the arrangement of sounds
3. Morphology is the study of word structure
4. Syntactics is the study of sentence structure
5. Semantics is the study of meanings in a sentence
6. Discourse is the study of syntactic and semantics on units of text longer than a sentence
7. Pragmatics is the study of language in communication

3.2 Context Awareness in Text Analysis

Different definitions of context were provided by linguists (Lichao, 2011). Widdowson (Widdowson, 1996) presented context as a schematic construction of the circumstances of language usage relevant to the meaning. Cook worked on the relationship between literature and discourse and used the context of texts as a form of global knowledge with a (1) broad definition or a (2) narrow definition (Cook, 1994). Lichao (Lichao, 2011) divided context into three categories:

1. Linguistic context refers to the context within the text. It consists of considering the relationship between words, phrases, sentences, and paragraphs. If we consider the word "bank" we need the sentence in which the word was mentioned to be able to properly assign the corresponding meaning. The study of time, place, and people related to happenings mentioned in a text form the deictic context element. Collocation of words falls into this context categorization. Collocation is the grouping of words with their context. For example: barked and dog, born and baby, blond and hair.

2. Situational Context is also known as the context of situation where the environment, time, place, participant of the text, and their relationship form the context. The activity domain or field of text, social relationships, and the mode of text communication are also part of this type of context.
3. Cultural Context as its name indicated considers the cultural background, customs, and past history of the language and the participant (writers or speakers of a discourse). Language is influenced by social factors, social status, gender, or age.

Depending on the discipline requirements, the context of texts in text analysis may play different roles:

- Eliminate Ambiguity in multiple levels: word sentence level and groups of sentences level.
- Improving coreference resolution and indicating referents which is generally used to replace noun phrases or adverbial phrases.
- Detecting Conversational Implicature or Intentions, ie; sarcasm, irony, insults, hurting, pain, caustic, humor, vulgarity, rhetorical questions, metaphors, …

There are two different conceptualizations of context and context used in NLP. The first conceptualization evokes the context of target words in their usage in a text. The second perspective of context is relative to knowledge extraction and the use of ontologies. A popular approach that enables the application of neural networks and machine learning algorithms is the representation of words, sentences, or documents in vectors, considering the "context" which in this case is the surrounding words (Kobayashi, 2018). The research field that provided this approach is Distributional Semantics based on the distributional structure of language theory proposed by Harris (Harris, 1954). Word2vect (Mikolov et al, 2013) and BERT (Devlin et al, 2020) are two models built based on de Distribution semantics. Word2vect was the first model released in 2013, it is unidirectional. In other words, in the example "I went to the bank to sit" and "I the bank to take some money", the word bank would have the same vector because the window considered is before the targeted. While BERT is the first Bidirectional Encoder Representation from Transformers, and the two representations of the word will be different.

The use of ontologies and ontology-based technics for information retrieval purposes was popular between 2000 and 2010 (Wimalasuriya et al, 2010). The use of predefined ontologies to orient and target the information, and the domain of the ontology would play the role of the context. It is that there is a bidirectional relationship between ontologies and natural language processing (Lenci, 2010). Ontologies can be used to orient knowledge extraction from text and NLP can help build and enrich ontologies. Lenci defined four major uses of contexts for onto-lexical knowledge extraction in NLP:

1. Semantic typing is used to characterize the semantic types of linguistic expressions
2. Identify semantic similarity and relatedness in which we try to pair words with similar meanings. In this context, the aim is to identify concepts that belong to the same logical type defined by Sommers (Sommers, 1963).
3. Enable inferences and inheritance of concepts within the same type
4. Argument structure which allows combining constraints of lexical items. Using predefined relationships, using ontologies enables the extraction of concepts and relationships between concepts based on lexical dependencies.

While distributional semantic approaches, such as BERT, are highly performant in machine learning tasks such as classification and annotation, a dependence on the pretraining dataset and the conceptualization of the group built the model. The models still require a need to structure and represent the information extracted while keeping track of the context of extraction. As for ontology-based context, the limit of this approach is the relativity of the knowledge extracted from the text and maintaining the context provided by the text. When using different texts, the categorization of concepts may not identify changes in concept definitions, evolution analysis of the context, and of the elements identified in the context. A need to track temporality and limit the inferences based on their context is identified. In the following section, we will present the context-awareness field, a research field dedicated to studying context and enabling systems to be aware of the context.

4 SITUATION CONTEXT RECOGNITION

Schilit et al. (Schilit, 1995) defined context-awareness as the ability of a system to adapt to a changing environment. In their use case, they worked on a system with mobile users, and the need was for the system to be able to detect the changing location and adapt to it. In their definition context was defined as the environment, the location of users, and the users. As the definition of context changed over time, the context-awareness definition also went through changes over time as the need for context-aware processing was needed. So, situation context can be recognized by identifying the situation object, the actors that provoked changes or made actions, the occurred event, the location of event, the event happening time, the field and domain of situation (Figure 2).

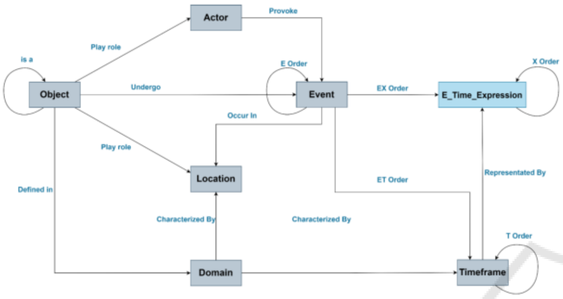
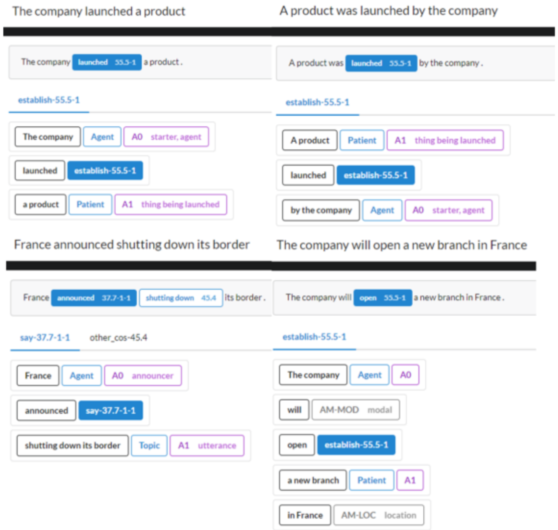


Figure 2: Main situation context elements (Matta et al, 2023).

The VerbNet parser will play a key role in identifying the agents in a sentence and their nature. VerbNet enables distinguishing between the agent ‘A0’ that is making the action for action verbs (Clark et al, 2021), (Brown et al, 2022). This takes into consideration the active and passive forms of the verb in a sentence. Note that the dimension actor in this approach will be represented by any entity that takes action in the text and will be the ‘A0’ provided by VerbNet. VerbNet also provides the part of the sentence that reflects the location and the time (Leseva et al, 2022).

Figure 3 provides an output of 4 different sentences. The 2 sentences on the top highlight the actor identification, regardless of the active or passive form of the verb, the parser detects the ‘company” as the “A0”, the starter of the action. As for the 2 sentences at the bottom, allows the distinction of roles the entity “France” is playing. In the first sentence, France is detected as an actor, therefore, instead of just being the country, it is also the political actor. While in the last sentence, France is the location of the event.

As for the objects, they represent any concept mentioned in the text and are extracted using hypernym/meronym relation extraction (Issa Alaa Aldine, 2022). The structural-based approach was used along with some of the Hearst Patterns (Roller et al, 2018) since they allow keeping track of the semantic extraction based on how it was mentioned in the text. It also enables the discovery of new relation and is not dependent on previously defined lexicons or context independent semantic relations. For instance, in the sentence “Domestic animals such as dogs and cats” we can identify the relation “animal/domestic animal” but also “domestic animal/dog” and “domestic animal/cat”. Ontologies are used to enhance inheritance among concepts within the same text and comparing objects from multiple text considering their context.


The company launched a product

The company launched 55.5-1 a product .

establish-55.5-1

The company Agent A0 starter, agent

launched establish-55.5-1

a product Patient A1 thing being launched

A product was launched by the company

A product was launched 55.5-1 by the company .

establish-55.5-1

A product Patient A1 thing being launched

launched establish-55.5-1

by the company Agent A0 starter, agent

France announced shutting down its border

France announced 37.7-1-1 shutting down 45.4 its border .

say-37.7-1-1 other_cos-45.4

France Agent A0 announcer

announced say-37.7-1-1

shutting down its border Topic A1 utterance

The company will open a new branch in France

The company will open 55.5-1 a new branch in France .

establish-55.5-1

The company Agent A0

will AM-MOD modal

open establish-55.5-1

a new branch Patient A1

in France AM-LOC location

Figure 3: Example of Using VerbNet Parser in CATKoRD platform (23), (24).

The CAToRD (Matta et al, 2023) has been developed based in these principles to identify situation context elements from text.

5 CULTURAL CONTEXT RECOGNITION

One dimension of cultural context can be related to the social evolution of a culture by analyzing the literature heritage of a civilization. When observing linguistics actors on literature text analysis, several dimensions can be identified:

- Text title type recognition that leads to text style identification.
- Text Blocs identification that emphasizes the organization of the document.
- Authors and references identification that help on literature type selection.
- Language Analysis to identify linguistics forms related to literature type.

As first step of the methodology, we want to define for cultural context recognition, a linguistic expert has been observed when analyzing two types of text from: French literature one and public scientific newspaper. These first results will help us to determine main aspects to consider in cultural context. That can be the foundations of cultural context ontology and NLP dedicated algorithms definition.

For instance, analyzing the "The Wolf and the Lamb" text leads to:

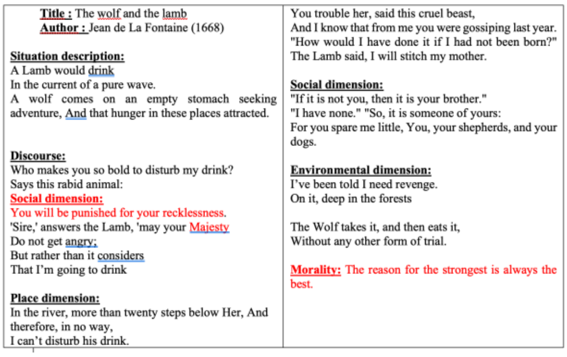
- Title of the text concerns two animals that have different characteristics.
- Text is decomposed on three blocs:
  1. An introduction that introduces the situation of two animals
  2. Discourses between two animals
  3. A conclusion on one sentence that emphasizes the text morality.
- Author: "La Fontaine", with reference, the title of the book: "Les Fables De La Fontaine" that leads to recognize the style of the literature; critics and morality documents.
- Linguistic analysis that helps to identify a conflict between a strong and weak characters.

These aspects push the linguistic analyst to isolate sentences that emphasizes this conflict (0):

- Space dimensions: The Wolf is higher than the lamb on the riverside.
- Social dimensions: social positions of the Wolf and the lamb in the society. "Magesty", "you and your family disorder my life" …
- Environmental dimensions: division of earth properties. "For you spare me little, You, your shepherds, and your dogs." …

This analysis put on a progress in conflict expression: from simple one between two animals to deeper one related to the humanity control of the environment and its impact.

In this type of text analysis, classical NLP algorithms are not sufficient to detect these types of dimensions. Analyzing sentences cannot enhance documents analysis. Cultural related to literature types must be defined. Semantic representations can be used as references that guide supervised NLP algorithms to detect such type of aspects.



Title : The wolf and the lamb
Author : Jean de La Fontaine (1668)

Situation description:
A Lamb would drink
In the current of a pure wave.
A wolf comes on an empty stomach seeking adventure, And that hunger in these places attracted.

Discourse:
Who makes you so bold to disturb my drink?
Says this rabid animal:
Social dimension:
You will be punished for your recklessness.
'Sire,' answers the Lamb, 'may your Majesty
Do not get angry;
But rather than it considers
That I'm going to drink

Place dimension:
In the river, more than twenty steps below Her, And therefore, in no way,
I can't disturb his drink.

You trouble her, said this cruel beast,
And I know that from me you were gossiping last year.
"How would I have done it if I had not been born?"
The Lamb said, I will stitch my mother.

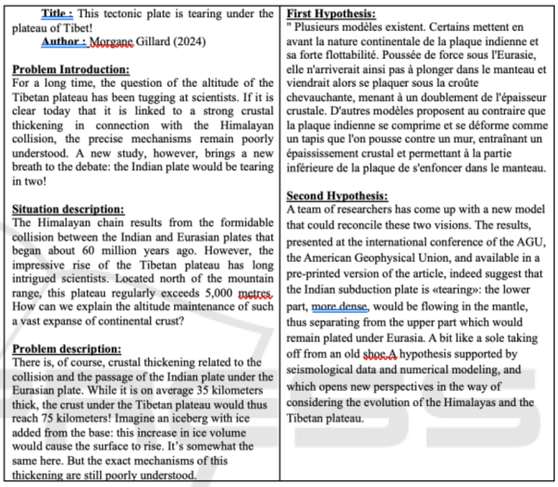
Social dimension:
"If it is not you, then it is your brother."
"I have none." "So, it is someone of yours:
For you spare me little, You, your shepherds, and your dogs.

Environmental dimension:
I've been told I need revenge.
On it, deep in the forests

The Wolf takes it, and then eats it,
Without any other form of trial.

Morality: The reason for the strongest is always the best.

Figure 4: Analysis of a Fable of La Fontaine[1].



Title : This tectonic plate is tearing under the plateau of Tibet!
Author : Morgane Gillard (2024)

Problem Introduction:
For a long time, the question of the altitude of the Tibetan plateau has been tugging at scientists. If it is clear today that it is linked to a strong crustal thickening in connection with the Himalayan collision, the precise mechanisms remain poorly understood. A new study, however, brings a new breath to the debate: the Indian plate would be tearing in two!

Situation description:
The Himalayan chain results from the formidable collision between the Indian and Eurasian plates that began about 60 million years ago. However, the impressive rise of the Tibetan plateau has long intrigued scientists. Located north of the mountain range, this plateau regularly exceeds 5,000 metres. How can we explain the altitude maintenance of such a vast expanse of continental crust?

Problem description:
There is, of course, crustal thickening related to the collision and the passage of the Indian plate under the Eurasian plate. While it is on average 35 kilometers thick, the crust under the Tibetan plateau would thus reach 75 kilometers! Imagine an iceberg with ice added from the base: this increase in ice volume would cause the surface to rise. It's somewhat the same here. But the exact mechanisms of this thickening are still poorly understood.

First Hypothesis:
" Plusieurs modèles existent. Certains mettent en avant la nature continentale de la plaque indienne et sa forte flottabilité. Poussée de force sous l'Eurasie, elle n'arriverait ainsi pas à plonger dans le manteau et viendrait alors se plaquer sous la croûte chevauchante, menant à un doublement de l'épaisseur crustale. D'autres modèles proposent au contraire que la plaque indienne se comprime et se déforme comme un tapis que l'on pousse contre un mur, entraînant un épaississement crustal et permettant à la partie inférieure de la plaque de s'enfoncer dans le manteau.

Second Hypothesis:
A team of researchers has come up with a new model that could reconcile these two visions. The results, presented at the international conference of the AGU, the American Geophysical Union, and available in a pre-printed version of the article, indeed suggest that the Indian subduction plate is «tearing»: the lower part, more dense, would be flowing in the mantle, thus separating from the upper part which would remain plated under Eurasia. A bit like a sole taking off from an old shoe.A hypothesis supported by seismological data and numerical modeling, and which opens new perspectives in the way of considering the evolution of the Himalayas and the Tibetan plateau.

Figure 5: Example of Analysis of a problem-solving scientific report[2].

Documents can be not only from the literature but belong to specific fields, activities, companies. For instance, technical reports are decomposed on several blocks, in which the problem is first described "*The question of the altitude of Tibetan problem…*", then observations are detailed before describing "*The Himalayan chain results from the formidable collision…*" how actors face the problem and give a solution. The title of these reports put on generally the encountered problem by presenting two hypotheses of "*floatability of the plate*" or "*separation of the Eurasian plates*". As same as, technical documents or contractual ones are presented related to reasoning schemas in which different blocks reflect actors' problem solving (0). Semantic representations and domain ontologies must be considered to emphasize

[1] https://www.poetica.fr/poeme-849/jean-de-la-fontaine-le-loup-et-agneau/

[2] https://www.futura-sciences.com/planete/actualites/tectonique-plaques- cette-plaque-tectonique-train-dechirer-sous-plateau-tibet-110914/

companies' cultural context and guide the NLP analysis of these type of documents.

6 CONCLUSION

Currently, NLP techniques tend to use LLM and Generative AI to analyze texts. But this type of techniques still be hard to consider specific context of activities which are necessary to emphasize semantics of a document. In fact, they ask to define several specific prompts (Feldman et al, 2023) and don't put on global techniques for this aim. In this paper, the importance to consider context in NLP algorithms has been shown based on our first studies on this domain. Firstly, some techniques to detect situation context has been mentioned and secondly, importance of cultural context to analyze documents are emphasized.

This paper presents our first study to detect cultural context. Two types of texts have been analyzed manually to identify important parts to consider in cultural context. We aim at studying cultural works to define a dedicated ontology. Then, CAToRD platform will be augmented by extending NLP algorithms that help to consider cultural context. Global rules will be then defined to be integrated in NLP applications and LLM algorithms.

REFERENCES

Adhikari A., Ram A., Tang R. , and Lin J., 'DocBERT: BERT for Document Classification', no. *arXiv:1904.08398. arXiv, Aug. 22, 2019. doi: 10.48550/arXiv.1904.08398.*

Adomavicius G., B. Mobasher B., F. Ricci B., and Tuzhilin A., 'Context-Aware Recommender Systems', *AI Magazine, vol. 32, pp. 67–80, Sep. 2011, doi: 10.1609/aimag.v32i3.2364.*

Bazire M. and P. Brézillon P., 'Understanding Context Before Using It', in *Modeling and Using Context*, A. Dey, B. Kokinov, D. Leake, and R. Turner, Eds., *in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2005, pp. 29–40. doi: 10.1007/11508373_3.*

Brown S.W., Bonn J., Kazeminejad G., Zaenen A., Pustejovsky J., and Palmer M., 'Semantic representations for nlp using verbnet and the generative lexicon', *Frontiers in artificial intelligence*, vol. 5, p. 821697, 2022.

Clark P., Dalvi B., and Tandon N., 'What Happened? Leveraging VerbNet to Predict the Effects of Actions *in Procedural Text', arXiv:1804.05435 (cs), Apr. 2018, Accessed: Oct. 20, 2021. (Online). Available: http://arxiv.org/abs/1804.05435*

Cook G., 'Discourse and Literature', *Oxford University Press, 1994, p. 24.*

Devlin J., Chang M.W., Lee K., and Toutanova K., 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding', arXiv:1810.04805 (cs), May 2019, Accessed: Nov. 02, 2020. (Online). Available: http://arxiv.org/abs/1810.04805

Du M., *Natural language processing system for business intelligence', p. 88, 2017.*

Feldman P., Foulds, J. R., & Pan, S. (2023). Trapping llm hallucinations using tagged context prompts. *arXiv preprint arXiv:2306.06085.*

Ghosh M., Roy M., Bandyopadhyay S., and Bandyopadhyay K., 'A tutorial review on *Text Mining Algorithms', Jun. 2012.*

Harris Z.S., 'Distributional structure', *Word*, vol. 10, pp. 146–162, 1954, doi: 10.1080/00437956.1954.11659520.

Issa Alaa Aldine A., 'Contributions to Hypernym Patterns Representation and Learning based on Dependency Parsing and Sequential Pattern Mining', *These de doctorat, Lorient, 2020. Accessed: Jun. 13, 2022. (Online). Available: http://www.theses.fr/2020LORIS575*

Kobayashi S., 'Contextual Augmentation: Data Augmentation by Words with Paradigmatic Relations', in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), New Orleans, Louisiana: Association for Computational Linguistics, Jun. 2018, pp. 452–457. doi: 10.18653/v1/N18-2072*

Lenci, A. The life cycle of knowledge. Ontology and the Lexicon. A Natural Language Processing Perspective. Cambridge University Press, Cambridge, UK, 241-257.2010

Leseva S. and Stoyanova, I. 'Linked Resources towards Enhancing the Conceptual Description of General Lexis Verbs Using Syntactic Information', in *Proceedings of the 5th International Conference on Computational Linguistics in Bulgaria (CLIB 2022)*, 2022.

Lichao S., 'The Role of Context in Discourse Analysis', *Journal of Language Teaching and Research*, vol. 1, Nov. 2010, doi: 10.4304/jltr.1.6.876-879.

Liddy E., 'Natural Language Processing', *School of Information Studies - Faculty Scholarship*, Jan. 2001, (Online). *Available: https://surface.syr.edu/istpub/63*

Mikolov T., Chen K., Corrado, G. and Dean J., 'Efficient Estimation of Word Representations *in Vector Space', no. arXiv:1301.3781. arXiv, Sep. 06, 2013. doi: 10.48550/arXiv.1301.3781.*

Matta N., Matta N., Giret E., and Declercq N., 'Enhancing Textual Knowledge Discovery using a Context-Awareness Approach', in *2021 International Conference on Computational Science and Computational Intelligence (CSCI)*, Dec. 2021, pp. 233–237. doi: 10.1109/CSCI54926.2021.00071.

Matta N., Matta N., Marcante A., and Declercq N., 'CATKoDR: Hybrid Context-Awareness Model

Architecture for Natural Language Processing', 2023, *Accessed: Apr. 08, 2024. (Online). Available: https://ieeesmc2023.org/abstract_files/SMC23_1543_FI.pdf*

Matta N., Matta N., Declercq N., and A. Marcante, 'Semantic Patterns to Structure TimeFrames in Text', *in INTELLI 2022, The Eleventh International Conference on Intelligent Systems and Applications, May 2022, pp. 16–23. Accessed: Sep. 13, 2022. (Online). Available: https://www.thinkmind.org/index.php?view=article&articleid=intelli_2022_1_40_60016*

Matsumoto D., 'Culture, Context, and Behavior', *Journal of Personality*, vol. 75, no. 6, Art. no. 6, 2007, doi: *10.1111/j.1467-6494.2007.00476.x.*

Roller S., Kiela S. and Nickel D. (2018). Hearst patterns revisited: Automatic hypernym detection from large text corpora. *arXiv preprint arXiv:1806.03191*.

Schilit W.N., *A system architecture for context-aware mobile computing. Columbia University, 1995.*

Sommers F., 'Types and Ontology', *The Philosophical Review, vol. 72, no. 3, Art. no. 3, 1963, doi: 10.2307/2183167.*

Widdowson H.G., 'Linguistics', in *Linguistics*, OUP Oxford, 1996, p. 126.

Wimalasuriya C. and D. Dou D., 'Ontology-based information extraction: An Introduction and a survey of current approaches', *J. Information Science, vol. 36, pp. 306–323, May 2010, doi: 10.1177/0165551509360123*.